Análise Fatorial Exploratória

TUANY DE PAULA CASTRO

Finalidades e Aplicações

- ✓ **Objetivo:** Reduzir o número de variáveis de uma base de dados, identificando o padrão de correlação ou de covariância entre elas em termos de quantidades aleatórias latentes (não observáveis) chamadas fatores.

Quando usar?

- Quando se tem interesse em resumir as informações contidas em grandes conjuntos de variáveis.
- Quando se tem interesse em mensurar fenômenos que não podem ser diretamente observados (inteligência, depressão, auto-eficácia, satisfação no trabalho, personalidade, bem-estar psicológico, agressividade, etc).

Finalidades e Aplicações

Agrupamento de variáveis de acordo com suas correlações:

- ✓ Variáveis dentro de um mesmo grupo altamente correlacionadas;
- ✓ Variáveis de grupos distintos com baixa correlação.

Exemplo

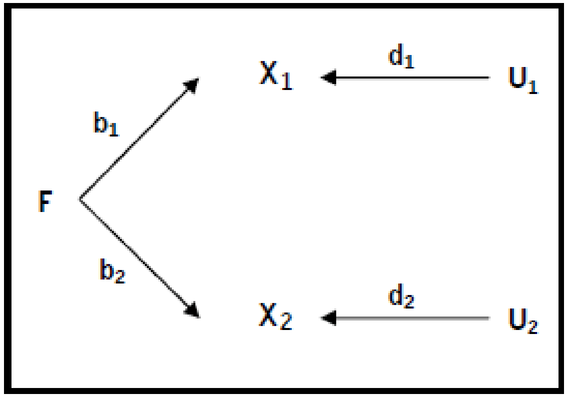
- As pontuações de alunos em disciplinas como Português, Inglês e Espanhol podem formar um fator latente de inteligência linguística;
- As pontuações de alunos em matemática, física e química podem formar um fator latente de inteligência lógica;
- As pontuações em esporte, dança, entre outros podem formar um fator latente de inteligência física-motora.

Finalidades e Aplicações

O problema consiste em obter combinações lineares das variáveis:

- Estas combinações lineares são chamadas fatores;
- Os fatores são variáveis latentes que dão origem às variáveis reais;
- Os fatores também variam entre os indivíduos, porém não podem ser medidos ou observados;
- A existência dessas variáveis hipotéticas é uma questão em aberto.

Figura 1
Modelo das vias para duas variáveis, modelo de um fator comum

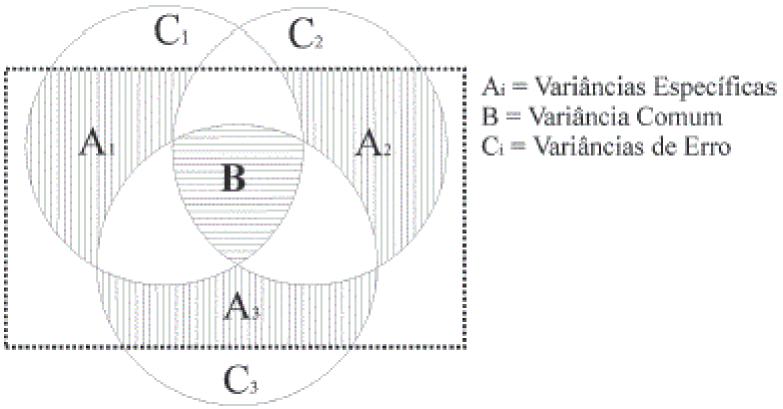


Fonte: Asher, 1983

Finalidades e Aplicações

Os métodos de redução de dados assumem que a variância de um fator ou componente é composto por três aspectos:

- **Variância específica:** variância de uma variável, não compartilhada com as demais;
- **Variância comum:** compartilhada entre todos os itens que compõem o fator ou componente;
- **Variância de erro:** parcela da variância não explicada pelo componente ou fator.



Finalidades e Aplicações

Diferença entre Análise Fatorial e Análise de Componentes Principais:

	Análise de Componentes Principais	Análise Fatorial
Variância	Não diferencia as variâncias comum e específica	Considera apenas a variância comum.
Índices	Representam a variância comum e a específica.	Representam apenas a variância comum.
Vantagens	Variabilidade explicada mais elevada.	Mais precisão na compreensão dos fatores .

Finalidades e Aplicações

Diferença entre Análise Fatorial e Análise de Componentes Principais:

“ Se você estiver interessado numa solução teórica não contaminada por variabilidade de erro, a análise fatorial deve ser sua escolha. Se você quiser simplesmente um resumo empírico do conjunto de dados, a análise de componentes principais é uma escolha melhor.” (Tabachinick e Fidell, 2007)

“A análise de componentes principais é em geral preferida para fins de redução de dados, enquanto a análise fatorial é em geral preferida quando o objetivo da pesquisa é detectar a estrutura dos dados da modelagem casual” (Garson, 2009)

História

No início do século XX, Karl Pearson, Charles Spearman e outros tentavam definir e medir a inteligência. Analisando pontuações de crianças em escolas, Spearman descobriu que a habilidade mental molda o desempenho cognitivo.

Devido a esta associação precoce com construções com a inteligência, a análise fatorial foi nutrida e desenvolvida principalmente por cientistas interessados em psicometria.

O advento de computadores de alta velocidade tem gerado um interesse renovado nos aspectos teóricos e computacionais da análise fatorial.

Raciocínio estrutural

Procedimento	O que deve ser observado
Verificar a adequabilidade da base de dados.	Nível de mensuração das variáveis, tamanho da amostra, razão entre o número de casos e a quantidade de variáveis e o padrão de correlação entre as variáveis.
Determinar a técnica de extração e o número de fatores a serem extraídos.	O tipo de extração (componentes principais, fatores principais, máxima verossimilhança, mínimos quadrados, mínimos quadrados ponderados)
Decidir o tipo de rotação dos fatores.	Se for ortogonal (varimax, quartimax, equamax), se for oblíqua

Critérios para o uso

Devem ser observados:

- Nível de mensuração
- Tamanho da amostra e razão entre número de casos e quantidade de variáveis
- Padrão de correlação entre as variáveis

Nível de mensuração

- ✓ Apenas variáveis discretas ou contínuas;
- ✓ Caso seja necessário, inclusão de categóricas como *dummies;*
- ✓ Algumas variáveis nunca devem ser incluídas por ser improvável a influência de fatores em sua variação (sexo e cor, por exemplo).

Critérios para o uso

Tamanho da amostra e razão entre número de casos e quantidade de variáveis

- ✓ Amostra superior a 50 observações;
- ✓ Aconselhável pelo menos 100 casos para resultados mais robustos;
- ✓ Razão mínima de 5 casos para uma variável.

Critérios para o uso

Padrão de correlação

Observar se a matriz de dados é passível de fatoração:

- **Critério de Kaiser-Meyer-Olkin (KMO)**
- **Teste de esfericidade de Bartlett**

Observação: Se a matriz for passível de fatoração, deve-se **avaliar se os dados seguem Distribuição Normal** para que possa ser escolhido um método de extração apropriado (por exemplo, máxima verossimilhança, principais eixos fatoriais, mínimos quadrados generalizados, mínimos quadrados não ponderados, fatoração alfa).

Critérios para o uso

- **Critério de Kaiser-Meyer-Olkin (KMO)**

  - Sugere a proporção da variância dos itens que pode estar sendo explicada por uma variável latente;
  - Varia de zero a um;
  - Valores iguais ou próximos de zero indicam baixa explicação pelas variáveis latentes, ou seja, a análise fatorial é inapropriada;
  - Regra para interpretação:

$$\begin{cases} KMO < 0{,}5: inaceitáveis \\ 0{,}5 \leq KMO < 0{,}7: medíocres \\ 0{,}7 \leq KMO < 0{,}8: bons \\ KMO \geq 0{,}8: ótimos \end{cases}$$

Critérios para o uso

- **Teste de esfericidade de Bartlett**

  - Avalia em que medida as variáveis não têm correlações entre si;
  - Valores do teste com nível de significância $p < 0{,}05$ indicam que a matriz de dados é fatorável, rejeitado a hipótese de que as variáveis são não correlacionadas.

Critérios para o uso

Nível de mensuração	Variáveis discretas e contínuas
Amostra	Amostras mínimas de 50 casos; razão mínima de 5 casos para uma variável.
Correlação	Maior parte dos coeficientes de correlação deve apresentar valores acima de 0,30.
KMO	Quanto maior, melhor, tendo 0,50 como o patamar mínimo de adequabilidade (Hair t al, 2006).
BTS	$p < 0,05$

Exemplo

Considere um conjunto de dados obtido a partir de um questionário sobre satisfação da vida aplicado a 100 adultos. O questionário contém 10 itens que são projetados para medir satisfação no trabalho, satisfação com *hobbies*, satisfação em casa e satisfação geral em outras áreas da vida. O objetivo é conhecer os fatores "por trás" desses diferentes domínios de satisfação e seus significados.

Exemplo

Work1	Work2	Work3	Hobby1	Hobby2	Home1	Home2	Home3	Miscel1	Miscel2
105,126	101,659	115,060	100,998	95,184	100,281	101,667	85,553	104,035	110,278
77,049	72,933	77,485	72,744	61,563	93,854	95,392	88,609	70,115	72,000
86,017	82,206	78,889	77,951	91,705	86,773	108,070	93,348	86,021	70,688
91,425	106,107	95,640	90,901	111,466	100,248	86,080	93,822	101,224	82,665
113,714	92,029	99,079	79,277	98,416	104,013	83,271	69,621	82,820	70,022
86,606	87,817	67,663	93,662	77,997	99,822	97,275	108,622	91,400	79,776
95,067	94,505	98,081	94,513	97,422	93,694	99,181	96,398	90,732	86,707
113,500	104,607	105,572	101,008	102,275	87,427	96,664	86,577	93,057	112,702
104,549	97,299	94,074	88,538	98,112	97,785	99,585	99,761	99,399	105,908
104,635	97,908	85,823	82,486	90,447	104,688	95,076	99,695	77,630	62,031
102,064	87,010	94,687	79,203	68,482	78,995	86,430	78,822	89,729	87,553
109,428	94,937	104,396	119,293	112,988	122,931	114,816	102,109	116,339	101,363
89,994	77,392	100,771	97,026	111,107	107,660	116,858	114,730	88,715	76,495
89,983	91,259	89,216	81,250	100,519	94,902	96,723	94,411	88,419	89,670

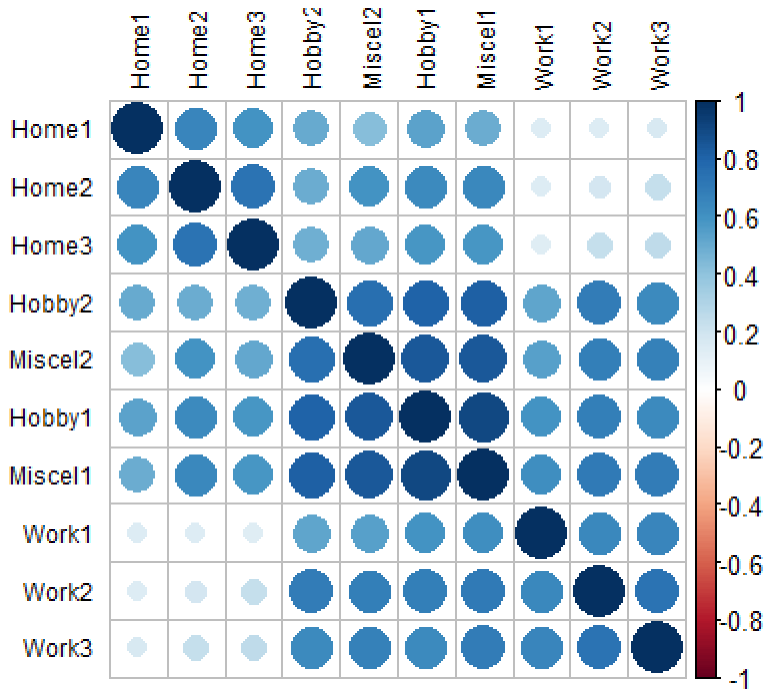
Exemplo

Critérios para o uso:

- **Nível de mensuração:** todas as variáveis são contínuas (escores);
- **Amostra:** 100 casos e razão de 10 casos para cada variável;
- **KMO = 0,91:** ótimo;
- **Teste de esfericidade de Bartlett:** valor-p = 0% < 5%, então rejeitamos a hipótese de variáveis não correlacionadas.

É necessário agora avaliar as correlações.

Exemplo



A maioria das correlações são maiores do que 0,30, indicando que a análise Fatorial é adequada.

Métodos de extração de fatores

Considere $X_1, X_2, \dots, X_p$ variáveis observáveis. Os fatores $F_1, F_2, \dots, F_m$, com $m < p$, são tais que:

$$X_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + e_1$$

$$X_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + e_2$$

.

.

.

$$X_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + e_p$$

Em que e_i são os erros e a_{ij} são as cargas fatoriais.

Métodos de extração de fatores

No exemplo das inteligências, teríamos:

$$Português = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + e_1$$

$$Inglês = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + e_2$$

$$Espanhol = a_{31}F_1 + a_{32}F_2 + a_{33}F_3 + e_3$$

$$Matemática = a_{41}F_1 + a_{42}F_2 + a_{43}F_3 + e_4$$

$$Física = a_{51}F_1 + a_{52}F_2 + a_{53}F_3 + e_5$$

$$Química = a_{61}F_1 + a_{62}F_2 + a_{63}F_3 + e_6$$

$$Esporte = a_{71}F_1 + a_{72}F_2 + a_{73}F_3 + e_7$$

$$Dança = a_{81}F_1 + a_{82}F_2 + a_{83}F_3 + e_8$$

$$Artes = a_{91}F_1 + a_{92}F_2 + a_{93}F_3 + e_9$$

Com F_1 a inteligência linguística, F_2 a inteligência lógica e F_3 a inteligência física-motora.

Os valores 'a' são chamados pesos ou cargas fatoriais.

Métodos de extração de fatores

No exemplo, os dados formavam uma nuvem de pontos num espaço de dimensão 10. Com a Análise Fatorial, transportamos os pontos para um espaço de dimensão 4 e, portanto, mais fácil de interpretar.

Considerações:

- ✓ Os fatores e os erros são independentes;
- ✓ Os fatores tem esperança igual a zero.

Métodos de extração de fatores

Análise de componentes principais

- ✓ Método amplamente utilizado para a extração do fator;
- ✓ Os pesos são calculados para extrair a variância máxima possível, continuando a fatoração sucessiva até que não haja mais variância significativa.

Análise fatorial canônica

- ✓ Também chamada de fatoração canônica de Rao;
- ✓ Método diferente de computação do mesmo modelo da Análise de Componentes Principais (método do eixo principal);
- ✓ Busca fatores que apresentam a maior correlação canônica com as variáveis observadas.

Métodos de extração de fatores

Análise Fatorial Comum ou Principais Eixos Fatoriais

- ✓Busca o menor número de fatores que podem explicar a variância comum (correlação) de um conjunto de variáveis;
- ✓ Comumente oferece bons resultados para amostras que apresentam distribuição não-normal.

Fatoração de imagens

- ✓ Baseia-se na matriz de correlação das variáveis preditas em vez das variáveis reais, onde cada variável é predita dos outros usando regressão múltipla.

Métodos de extração de fatores

Fatoração Alfa

- ✓ Baseia-se na maximização da confiabilidade dos fatores, assumindo que as variáveis são amostradas aleatoriamente de um universo de variáveis.

Modelo de Regressão Fatorial

- ✓ Modelo combinatório do modelo fatorial e do modelo de regressão, com fatores parcialmente conhecidos.

Modelo de Máxima Verossimilhança

- ✓ Supõe conhecido o número de fatores;
- ✓ O modelo estimará as cargas e as comunalidades que maximizam a probabilidade da matriz de correlação observada;
- ✓ Comumente oferece bons resultados para amostras que apresentam distribuição normal.

Métodos de extração de fatores

Após a extração dos fatores, algumas medidas são obtidas e devem ser interpretadas:

Cargas fatoriais

- ✓ Podem ser interpretadas como as correlações entre os fatores e as variáveis;
- ✓ A soma dos quadrados das cargas fatoriais dividido pela variância total é o percentual da variância explicada pelo fator;
- ✓ O sinal da carga indica apenas que variáveis com cargas opostas no mesmo fator se relacionam com esse fator de maneiras opostas;
- ✓ Pode-se multiplicar todas as cargas de um fator por -1 e os resultados não serão afetados.

Métodos de extração de fatores

Autovalor (*eigenvalue*)

- ✓ O autovalor de determinado fator indica quanto da variância das variáveis observadas é explicada por esse fator;
- ✓ Um fator com baixo autovalor contribui pouco para a explicação das variáveis e pode ser ignorado por ser considerado redundante.

Comunalidade

- ✓ A soma dos quadrados de todas as cargas fatoriais de todos os fatores para determinada variável observada é a variabilidade dessa variável que é representada pelos fatores, chamada comunalidade;
- ✓ Mede o percentual de variância de uma determinada variável observada explicada pelos fatores em conjunto;
- ✓ Pode ser interpretada como a confiabilidade do indicador.

Métodos de extração de fatores

Coeficientes dos escores fatoriais

- ✓ Utilizados para o cálculo dos escores fatoriais.

Escores fatoriais

- ✓ Valores "observados" dos fatores obtidos para cada unidade amostral;
- ✓ Útil para análises adicionais dos fatores.

Métodos de extração de fatores

O objetivo da extração de fatores é **determinar a quantidade de fatores que melhor representa o padrão de correlação entre as variáveis observadas**. Deve-se tem em conta dois princípios:

- **Parcimônia**: explicar as correlações entre as variáveis observadas utilizando o menor número possível de fatores;
- **Interpretabilidade**: fatores com significado no contexto estudado, com coerência lógica.

Métodos de extração de fatores

Após escolhido o método de extração dos fatores, deve-se definir quantos fatores serão extraídos. Não há um critério consensual, porém há alguns métodos para auxiliar na definição do número de fatores:

Critério de Kaiser

✓ Devem ser extraídos apenas os fatores com valor de *eigenvalue* (autovalor) acima de um.

Scree plot

✓ Analisar graficamente a dispersão do número de fatores até que a curva da variância individual de cada fator se torne horizontal ou sofra uma queda abrupta. Em ambas as situações, isso indica que muita variância foi perdida e, por isso, deve-se parar de extrair fatores.

Critério da variância acumulada

✓ Extrair fatores até que se alcance 60% da variância explicada.

Exemplo

Utilizando o método de extração de Análise de Componentes Principais, temos abaixo os autovalores e variâncias explicadas pelos fatores.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Autovalor	6,12	1,80	0,47	0,41	0,31	0,29	0,19	0,17	0,14	0,08
Proporção da variância	0,61	0,18	0,05	0,04	0,03	0,03	0,02	0,02	0,01	0,01
Proporção acumulada	0,61	0,79	0,84	0,88	0,91	0,94	0,96	0,98	0,99	1,00

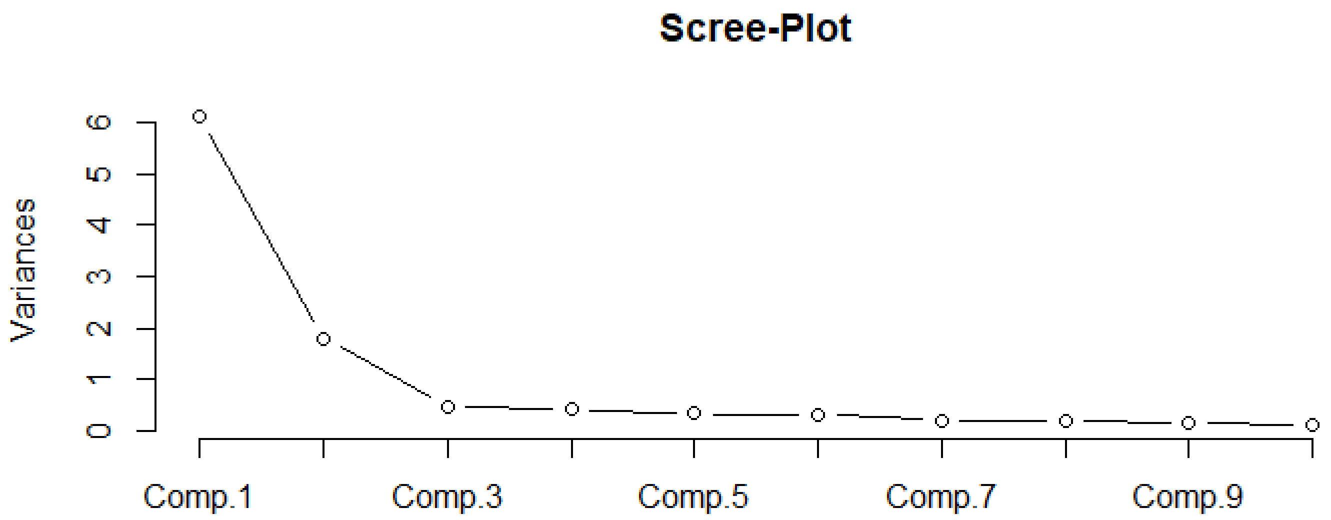
Exemplo

Decidindo o número de fatores:

- **Critério de Kaiser:** dois fatores, pois apenas eles têm autovalores maiores do que 1;
- **Critério da variância acumulada:** dois fatores, pois com eles já garantimos mais de 60% da variância explicada.

Agora vamos analisar o Scree Plot.

Exemplo



O ponto de corte é aquele em que a curva se torna horizontal, pois pouca informação é adicionada com a extração dos fatores adicionais. Aqui, esse ponto pode ser o 2 ou o 3, sendo indicado a análise dos dois casos para verificar aquele que trará padrões mais interpretáveis.

Exemplo

Para analisar as cargas fatoriais:

```
> loadings(fit) # cargas fatoriais

Loadings:
        Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9 Comp.10
Work1   -0.264 -0.383 -0.439  0.687         0.235 -0.116         0.215        
Work2   -0.306 -0.369  0.115 -0.332 -0.161  0.318  0.533 -0.395  0.279        
Work3   -0.301 -0.340  0.152        -0.364 -0.779                      -0.133 
Hobby1  -0.381                       0.214  0.173               -0.655 -0.589 
Hobby2  -0.354        -0.145 -0.508         0.168 -0.707         0.239        
Home1   -0.233  0.451 -0.714 -0.180 -0.200 -0.212  0.328                      
Home2   -0.271  0.461  0.183  0.250  0.399 -0.192        -0.487  0.391 -0.165 
Home3   -0.259  0.428  0.391  0.239 -0.644  0.295         0.194               
Miscel1 -0.385                       0.136               -0.232 -0.422  0.766 
Miscel2 -0.364         0.221         0.402         0.272  0.709  0.236  0.104 
```

Exemplo

Cargas fatoriais com o método de extração de Eixos Principais:

```
> # Método de extração: eixos principais - sem rotação
> fit_ep <- fa(dados, 2, rotate="none", fm = "pa")
> fit_ep
Factor Analysis using method =  pa
Call: fa(r = dados, nfactors = 2, rotate = "none", fm = "pa")
Standardized loadings (pattern matrix) based upon correlation matrix
         PA1   PA2   h2    u2 com
Work1   0.62 -0.43 0.57 0.432 1.8
Work2   0.74 -0.48 0.79 0.214 1.7
Work3   0.72 -0.42 0.69 0.306 1.6
Hobby1  0.94  0.02 0.89 0.111 1.0
Hobby2  0.85 -0.05 0.73 0.269 1.0
Home1   0.54  0.50 0.55 0.452 2.0
Home2   0.66  0.62 0.82 0.179 2.0
Home3   0.61  0.51 0.63 0.366 1.9
Miscel1 0.96 -0.01 0.92 0.085 1.0
Miscel2 0.89 -0.04 0.79 0.213 1.0

                       PA1  PA2
SS loadings           5.89 1.48
Proportion Var        0.59 0.15
Cumulative Var        0.59 0.74
Proportion Explained  0.80 0.20
Cumulative Proportion 0.80 1.00
```

Exercícios

1) Uma companhia de software está para lançar um novo software de processamento de texto. Em um estudo piloto, foi testada a importância de diversas variáveis por meio de questionamentos a possíveis usuários. As variáveis usadas no estudo foram:

- Aparência da interface;
- Preço;
- Compatibilidade com todos os tipos de impressoras;
- Qualidade do manual do usuário;
- Facilidade de definição de estilos de usuários;
- Número de estilos pré-definidos;
- Facilidade com ajudas;

Exercícios

- Compatibilidade com outros softwares;
- Qualidade da saída;
- Uso de codificação em cores nos estilos;
- Portabilidade entre plataformas de máquinas;
- Facilidade de customização;
- Número de cliques necessários;
- Possibilidade de importar documentos;
- Número de facilidades;
- Facilidade de uso.

Exercícios

Agora o interesse é decidir quais fatores são associados às variáveis para pontuá-los na propaganda do software. Sem utilizar técnicas estatísticas, você poderia identificar um ou mais fatores da lista de variáveis? Você poderia designar nomes a esses fatores?

2) Suponha que, em uma análise fatorial com um número grande de variáveis, 15 fatores foram extraídos dos dados com os seguintes autovalores: 8.3, 6.5, 5.4, 4.2, 3.0, 2.0, 1.2, 0.8, 0.6, 0.6, 0.5, 0.5, 0.4, 0.3 e 0.25.

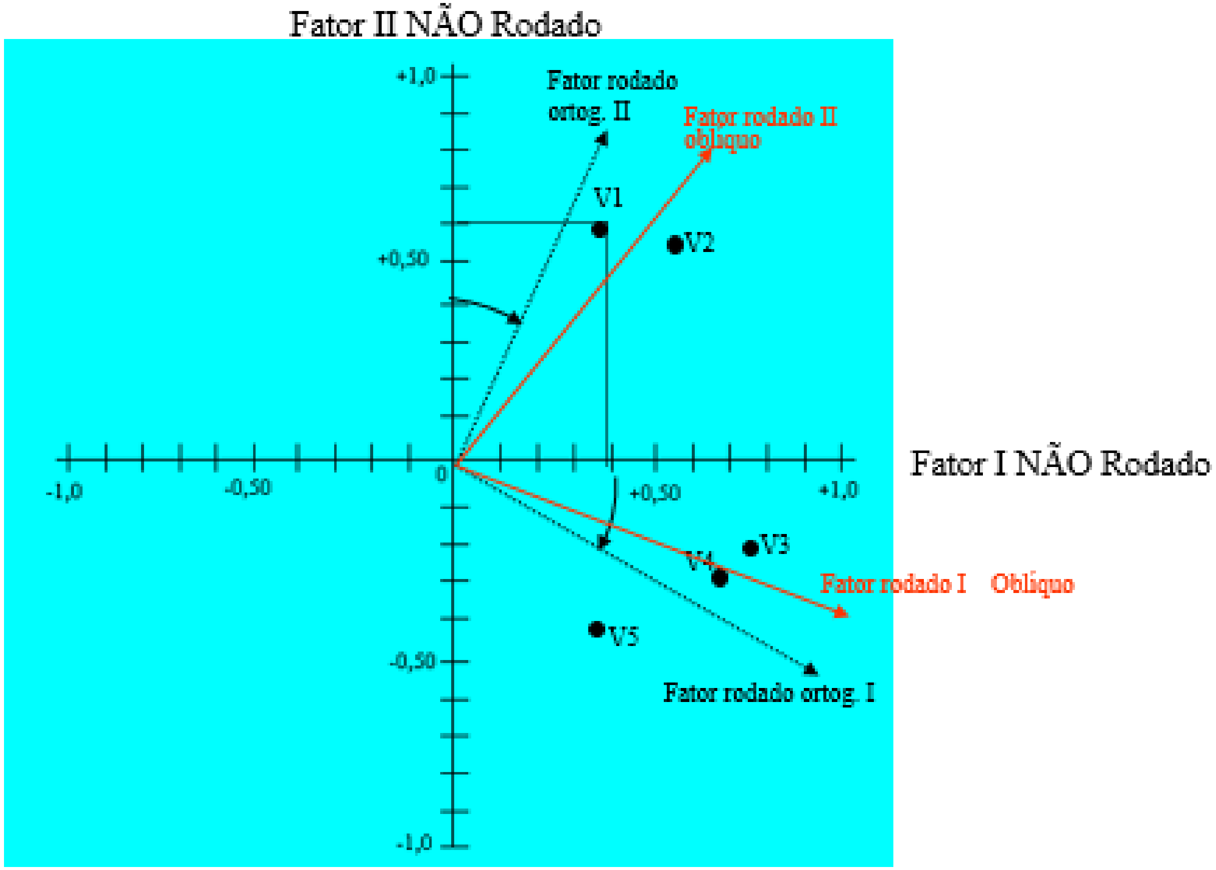
A) Faça o Scree-Plot para o montante de variância explicada pelos sucessivos fatores. Quantos fatores você incluiria na análise?

B) Agora utilize a regra de Kaiser. Quantos fatores você analisaria?

Métodos de rotação de fatores

- É um método matemático que rotaciona os eixos no espaço geométrico;
- Esse método torna mais fácil determinar quais variáveis são carregadas em quais componentes;
- **Objetivo**: tornar o resultado empírico encontrado mais facilmente interpretável, conservando as suas propriedades estatísticas.

Métodos de rotação de fatores



Métodos de rotação de fatores

Existem dois tipos de rotação:

- Ortogonal: supõe fatores independentes. Mais fácil interpretação:
  - Varimax
  - Quartimax
  - Equamax

- Oblíqua: fatores podem ser correlacionados, porém mais difícil interpretação:
  - Oblimin
  - Promax

Métodos de rotação de fatores

Varimax

- Opção de rotação mais comum;
- Procura minimizar o número de variáveis que apresentam altas cargas em cada fator.

Quartimax

- Procura minimizar o número de fatores necessários para explicar cada variável;
- Geralmente gera um fator geral sobre o qual a maioria das variáveis são carregadas em um grau alto ou médio.

Equamax

- Compromisso entre os critérios Varimax e Quartimax.

Métodos de rotação de fatores

Oblimin

- ✓ Método oblíquo padrão (fatores podem ser correlacionados);
- ✓ Autovalores mais altos, mas interpretabilidade diminuída.

Promax

- ✓ Método oblíquo alternativo, computacionalmente mais rápido.

Exemplo

Extração por Componentes Principais com Rotação Varimax.

```
> ################## Componentes Principais com Rotação Varimax ###########
#
> fit_acp_var <- principal(na.omit(dados), nfactors=2, rotate="varimax")
> fit_acp_var # print results
Principal Components Analysis
Call: principal(r = na.omit(dados), nfactors = 2, rotate = "varimax")
Standardized loadings (pattern matrix) based upon correlation matrix
         RC1   RC2   h2    u2 com
Work1   0.83 -0.01 0.69 0.310 1.0
Work2   0.90  0.07 0.82 0.182 1.0
Work3   0.87  0.10 0.76 0.235 1.0
Hobby1  0.73  0.59 0.89 0.113 1.9
Hobby2  0.72  0.50 0.77 0.231 1.8
Home1   0.08  0.83 0.70 0.302 1.0
Home2   0.15  0.90 0.83 0.167 1.1
Home3   0.15  0.85 0.74 0.259 1.1
Miscel1 0.76  0.57 0.91 0.094 1.9
Miscel2 0.74  0.51 0.81 0.187 1.8

                       RC1  RC2
SS loadings           4.50 3.42
Proportion Var        0.45 0.34
Cumulative Var        0.45 0.79
Proportion Explained  0.57 0.43
Cumulative Proportion 0.57 1.00

Mean item complexity =  1.4
Test of the hypothesis that 2 components are sufficient.

The root mean square of the residuals (RMSR) is  0.05
 with the empirical chi square  18.96  with prob <  0.84

Fit based upon off diagonal values = 0.99
> |
```

Para o primeiro fator, observam-se cargas altas (maiores do que 0,40) em todas as variáveis, exceto as de satisfação com a casa. Cargas mais substanciais para as questões relacionadas a trabalho.
Já no segundo fator, aparecem cargas substanciais para as questões de casa e apenas as variáveis relacionadas à satisfação no trabalho não são altas.

Exemplo

Comunalidades com a extração por componentes principais e rotação Varimax.

```
Console F:/UNICID/Análise Fatorial/1-AFE/Dados/
> # comunalidades
> fit_acp_var$communality
    Work1     Work2     Work3    Hobby1    Hobby2     Home1     Home2 
0.6903094 0.8178119 0.7646331 0.8871432 0.7693687 0.6978478 0.8325067 
    Home3   Miscel1   Miscel2 
0.7409493 0.9055663 0.8129168 
```

Observamos que todas as variáveis tiveram alto percentual de explicação (a partir de 0,69).

Exemplo

Podemos obter os escores dos indivíduos da amostra:

```
> fit_acp_var$scores
           RC1          RC2
1   0.78266580 -0.586754637
2  -1.95221890 -0.459399532
3  -1.31571876 -0.156484386
4   0.19038511 -0.705475089
5   0.10877673 -1.639832620
6  -1.43111429  0.399912145
7  -0.18783315 -0.397280175
8   0.96991038 -1.114955219
9   0.03672154 -0.205285260
10 -0.70030235 -0.421943339
11 -0.15790083 -1.760799381
12  0.21668865  1.194658518
13 -1.11421016  1.228552578
14 -0.56797346 -0.384698028
15  0.17537233 -0.086497407
16 -0.59302232  1.264803719
17 -2.52034489 -0.939499685
18  0.54940027 -1.104041795
19 -0.27812387 -0.006718397
20  0.95053189 -0.480087765
```

Exemplo

Cargas fatoriais com rotação varimax para o método de Eixos Principais:

```
Console F:/UNICID/Análise Fatorial/1-AFE/Dados/
> ?fa
> fit_ep <- fa(dados, 2, rotate="varimax", fm = "pa")
> fit_ep
Factor Analysis using method =  pa
Call: fa(r = dados, nfactors = 2, rotate = "varimax", fm = "pa")
Standardized loadings (pattern matrix) based upon correlation matrix
         PA1  PA2   h2    u2 com
Work1   0.75 0.04 0.57 0.432 1.0
Work2   0.88 0.08 0.79 0.214 1.0
Work3   0.82 0.12 0.69 0.306 1.0
Hobby1  0.73 0.60 0.89 0.111 1.9
Hobby2  0.70 0.49 0.73 0.269 1.8
Home1   0.12 0.73 0.55 0.452 1.1
Home2   0.14 0.89 0.82 0.179 1.0
Home3   0.17 0.78 0.63 0.366 1.1
Miscel1 0.76 0.58 0.92 0.085 1.9
Miscel2 0.72 0.51 0.79 0.213 1.8

                       PA1  PA2
SS loadings           4.20 3.17
Proportion Var        0.42 0.32
Cumulative Var        0.42 0.74
Proportion Explained  0.57 0.43
Cumulative Proportion 0.57 1.00
```

Resultados próximos ao de Componentes Principais.

Exercícios

3)

A) Calcule os autovalores, o percentual de variância explicada e acumulada e as comunalidades para o seguinte resultado de uma análise fatorial:

Variável	Fator 1	Fator 2
Sabor	0,56	0,82
Boa compra	0,78	-0,52
Aroma	0,65	0,75
Ideal como lanche	0,94	-0,10
Boa fonte de energia	0,80	-0,54

Exercícios

B) Construa o gráfico para o resultado anterior e veja como as variáveis estão localizadas nos fatores.

C) Calcule agora os autovalores, o percentual de variância explicada e acumulada e as comunalidades para os fatores anteriores rotacionados:

Variável	Fator 1	Fator 2
Sabor	0,02	0,99
Boa compra	0,94	-0,01
Aroma	0,13	0,98
Ideal como lanche	0,84	0,43
Boa fonte de energia	0,97	-0,02

Exercícios

D) Construa o gráfico para do item C e veja a diferença com relação ao gráfico do item B.

4) Considere o conjunto *dados_personalidade_32.csv* obtido da aplicação de um teste para 240 indivíduos sobre auto-avaliação de 32 traços de personalidade.

- (a) Avalie e comente a adequabilidade da Análise Fatorial.
- (b) Utilizando método de extração de Componentes Principais, defina quantos fatores você usaria no modelo. Justifique.
- (c) Comente sobre as comunalidades.
- (d) Com o método de rotação Varimax, interprete os fatores obtidos.
- (e) Repita os itens anteriores utilizando o método de extração Eixos Principais.

Referências

- BISQUERRA, R; CASTELLA, J.; VILLEGAS, F. Introdução à estatística: enfoque informático com o pacote estatístico SPSS. Porto Alegre: Artmed, 2007.
- DANCEY, Cristine P; REIDY, John. Estatística sem Matemática Para Psicologia. 3 edição. Porto Alegre: Artmed, 2006.
- HAIR, J.; ANDERSON, R.; BLACK, W. Análise multivariada de dados. 6 ed. Reimp. Porto Alegre: Bookman, 2009.
- JOHNSON, R. and WICHERN, D. Applied Multivariate Statistical Analysis. Sixth edition, Wisconsin, Pearson, 2007.